Does incongruency of lexico-semantic and prosodic information cause discernable cognitive conflict?

Running Head: Emotional Incongruity in Speech

Rachel L. C. Mitchell

School of Psychology and Clinical Language Sciences, University of Reading, U.K

Neuroscience and Emotion Section, Institute of Psychiatry, U.K.

Berkshire Healthcare NHS Trust, U.K.

Correspondence Address:

School of Psychology,

University of Reading,

Whiteknights Road,

Reading,

Berkshire.

RG6 6AL.

U.K.

Tel: +44 (0)118 378 7530

Fax: +44 (0)118 378 6715

Email: r.l.c.mitchell@rdg.ac.uk

Abstract

We are often required to interpret discordant emotional signals. Whilst equivalent cognitive paradigms cause noticeable conflict via their behavioural and psychophysiological effects, the same may not necessarily be true for discordant emotion. The skin conductance (SCR) and heart rate (HR) of participants were measured during a classic Stroop task and one in which the emotions conveyed by lexico-semantic content and prosody were congruent or incongruent. Participants' task was to identify the emotion conveyed by lexico-semantic content or prosody. No relationship was observed between HR and congruency. SCR was higher during incongruent conditions of the experimental task compared to congruent (ditto classic Stroop task), but no difference was observed when congruency effects during lexico-semantic emotion identification were compared to those during prosodic emotion identification. It is concluded that incongruence between lexico-semantic and prosodic emotion does cause notable cognitive conflict. Functional neuroanatomic implications are discussed.

Author Note

Name:

Dr. Rachel L. C. Mitchell

Affiliations:

School of Psychology & Clinical Language Sciences, University of Reading

Neuroscience & Emotion Section, Institute of Psychiatry

Research Psychologist, Berkshire Healthcare NHS Trust

Address:

School of Psychology & Clinical Language Sciences, University of Reading, Whiteknights Road, Reading, Berkshire, RG6 6AL, U.K.

Source of Research Support:

School of Psychology, Keele University, Staffordshire, U.K.

Introduction

Conflicting Emotional Signals

Everyday communications constantly require us to decipher emotional signals, to interact with other people in a successful and appropriate manner (Trauner, Ballantyne, Chase & Tallal, 1993), and to maintain social relationships (Carton, Kessler & Pape, 1999). To do this, we rely upon various cues, including facial emotion, gestures, (lexico-) semantic content, and prosody; communication and interpretation of emotion is a multi-determined behaviour (Borod, 1993). However, we are repeatedly faced with situations in which sources of emotional information contradict one other. Someone may describe a situation using words leading listeners to believe they were happy, but their facial expressions may convey a different emotion. Or a person may describe this situation using an angry tone of voice. In these scenarios the protagonist may wish to hide their true opinions or feelings. It is only in cases such as irony, banter and sarcasm that the incongruency itself possesses a particular, interpretable meaning.

Cognitive Conflict

In cognitive psychology, the Stroop task has become a classic paradigm with which to study the correlates of conflict and factors that modulate it. In the classic task participants are required to identify the ink colour in which colour words are printed. However, they find it difficult to suppress the more automatic process of reading. The occurrence of conflict at the cognitive level can be inferred from various measures, including certain patterns of behavioural and psychophysiological responses. At the behavioural level, participants are typically faster to name the ink colour for stimuli in which ink colour and colour word are congruent than for stimuli in which they are incongruent. (Macleod, 1991; Stroop, 1935). Skin conductance responses (SCR) are the rapid, transient changes in conductance in the surface of the skin that accompany changes in the autonomic innervation of eccrine sweat glands via the sympathetic nervous system. Increased cognitive and behavioural demands often lead to increased autonomic nervous system activity with consequential increases in sweat gland activity that are particularly noticeable in the fingers and palms of the hand due to their high concentration of sweat glands. When processing incongruent stimuli in the Stroop task, participants' behavioural reactions are usually accompanied by a noticeable increase in SCR (Tulen, Moleman, van Steenis & Boosma, 1989), indicating increased stress and arousal at the peripheral nervous system level (Hugdahl, 1995).

Detecting SCR Responses to Emotional Stimuli

When processing the emotional signals we encounter, individuals will display affective, behavioural, and facial reactions, and visceral reactions such as fluctuations in SCR (Lang, Greenwald, Bradley & Hamm, 1993). A multi-faceted response to emotions is intuitive when one considers the evidence for associations between these different systems. Indeed, James (1884) claims that emotions ARE the autonomic nervous system's response to emotion provoking stimuli. Furthermore, direct neural pathways exist between long-established emotional processing regions such as the amygdala and executive function regions such as the prefrontal cortex and anterior cingulate cortex (ACC), and these pathways are thought to enable goal-directed behavioural responses to cognitive tasks (Holland & Gallagher, 2004). These same brain regions are amongst those known to influence SCR (Critchley, 2002). Thus, SCRs have now been found to be affected by facial, pictorial, video and written emotion stimuli (Campos, Marcos & González, 1999; Gomez, Stahel & Danuser, 2004; Hubert & de Jong-Meyer, 1990; Lang et al., 1993; Meadows & Caplan, 1994; vanOyen Witvliet & Vrana, 1995). Despite the interest in interactions between emotions and psychophysiological responses, comparatively little is currently known about the psychophysiological response to auditory emotion.

Why Might Incongruent Emotional Information Not Cause Cognitive Conflict?

The key element of paradigms (such as the Stroop task) typically used to induce cognitive conflict is the inter-dimensional conflict of task stimuli. Since this feature is present in speech in which lexico-semantic and prosodic cues convey different emotions, a cursory consideration of this similarity might cause one to think that emotional incongruency in speech would similarly induce cognitive conflict. However, there are reasons to suspect that this might not necessarily be the case. In children, the ability to decode emotional signals develops at a surprisingly early age, with infants as young as 3 years old achieving near perfect performance on tests of facial and prosodic emotion recognition (Brosgole & Weisman, 1995). Often, when a task or ability is practiced repeatedly, a person becomes quite proficient, sometimes to the point of internalising the ability as a skill, and reducing the conscious effort needed to perform the ability again (Shanks & Johnstone, 1999). Interpersonal skills can similarly be improved by practice (McConnell, 2004). Thus if we learn the ability to decode emotional signals at an early age and continue to engage in social interactions requiring these skills, by adulthood an (healthy) individual should be highly skilled at performing this ability. As for the added demands posed by incongruent emotional signals, evidence from classic Stroop task paradigms (MacLeod & Dunbar, 1988; Melara , Rao & Tong, 2002) would suggest that the cognitive conflict posed by incongruous relevant and task irrelevant dimensions similarly diminishes with practice, although both the cognitive and emotional form, may require extended periods of practice.

Previous Applications of Related Tasks

Whilst paradigms incorporating conflict between lexico-semantic and prosodic emotion cues have been used in previous research (see Table 1), behavioural evidence of between-dimension congruency effects is inconsistent (see below). For this reason alone, reliance on behavioural performance measures may not provide a clear or reliable picture of whether lexico-semantic/prosodic incongruency causes cognitive conflict. Further, whereas we may mask or alter our true behavioural reactions via conscious choice, it is much more difficult to consciously choose to alter our psychophysiological reactions (Dollins , Cestaro & Pettit, 1998). They may therefore represent a more objective index of the response to cognitive conflict. Event-related potentials (ERP) have been used to index the central nervous system response to incongruency between lexico-semantic and prosodic cues, and the report of this application interpreted an exaggerated N400 response as evidence for a ‘Stroop’ type effect (Schirmer & Kotz, 2003). However, even in healthy populations, ERP responses are subject to broad inter-participant variability of peak latencies and amplitudes, the specificity of some ERP tests is questionable (Barrett, 2000), and others have raised doubts about the separability of endogenous and exogenous components of the ERP response (Gaillard, 1988). Another means of assessing the internal response to lexico-semantic/prosodic inongruency may therefore be prudent. In the current study discord between the emotions conveyed by lexico-semantic and prosodic cues were examined as they would occur in most everyday speech, i.e. in sentences. Previous applications of these paradigms have either examined the effects of congruency in relation to isolated words (Grimshaw, 1998; Kitayama & Ishii, 2002; Schirmer & Kotz, 2003), or in carrier sentences in which the lexico-semantic and prosodic information of interest is limited to the end word (Wurm & Vakoch, 1996; Wurm , Vakoch, Strasser, Calin-Jageman & Ross, 2001).

Table 1 about here

Hypotheses & Rationales

In the study outlined below, the following hypotheses were tested

(i) SCR when processing incongruent lexico-semantic and prosodic emotion would be higher than that when they were congruent.

See discussion above.

At the behavioural level, data (from paradigms in which lexico-semantic and prosodic information are varied in congruence) appears equivocal, with reports of supporting (Grimshaw, 1998; Schirmer & Kotz, 2003) and negative evidence (Wurm & Vakoch, 1996; Wurm et al., 2001).

(ii) The effect of congruency on SCR when responding to lexico-semantic emotion would be different to that when responding to prosodic emotion.

At the behavioural level, evidence is again contradictory. According to Kitayama & Ishii (2002), incongruent lexico-semantics and prosody only caused cognitive conflict when participants made judgements about prosody. Grimshaw (1998) also found that the effect of congruency was higher when making judgements about prosody than when judging lexico-semantic emotion, but the effect of congruency when making lexico-semantic judgements was still significant. In contrast, Schirmer & Kotz (2003) found that the congruency effect was stronger for lexico-semantic judgements than judgements about prosody.

(iii) The rise in SCR caused by incongruity of information from task-relevant and irrelevant stimulus dimensions (vs. congruity) would be greater for the classic Stroop task than for the auditory emotional interference task.

For the reasons posed above regarding practice at this social skill, it is hypothesised that if emotional incongruity does induce cognitive conflict, that the size of the effect will be notably smaller than for classic inducers of cognitive conflict.

Methods

This study was granted approval by the ethics committee of the School of Psychology, Keele University, U.K.

Participants

Forty-four psychiatrically normal participants (36 females, 8 males) with a mean age of 18.7 years ($\pm$1.3) were recruited from undergraduates in the department of psychology, Keele University and from local A-level students. Absence of psychiatric illness was established via self-report. To avoid confounding behavioural data, volunteers with colour blindness or hearing problems were excluded. To facilitate accurate psychophysiology, volunteers with high blood pressure or heart problems were excluded.

Experimental Paradigms

In this study, the effects of inter-dimensional incongruity in the auditory emotional interference paradigm were compared to that in a classic colour Stroop task. Inclusion of the classic colour Stroop task served two main purposes. First, it functioned as a control task, to rule out negative results for the auditory emotional interference paradigm being caused by faulty equipment or procedure, and secondly it functioned as a standard comparator, to gauge the size of any cognitive conflict induced by the auditory emotional interference paradigm. Administration order of the two tasks was counterbalanced across participants.

Classic colour Stroop task – Two series of 40 colour words (red, green, blue, brown) were presented in the centre of the screen (every 2 s, with a 0.5s gap), in different colour fonts (red, green, blue, brown) in Arial point 18. For the congruent condition, colour words were displayed (in a single block) in matching font colours, whereas in the incongruent condition colour words were displayed (again in a single block) in a conflicting colour (e.g. the word red displayed in blue). In each condition participants were instructed to indicate the font colour as quickly as they could, by pressing 1 for red, 2 for green, 3 for blue and 4 for brown. The order of the congruent and incongruent blocks was alternated between participants.

Auditory emotional interference task – Pre-recorded sentences were used for which the emotion (happy or sad) conveyed by lexico-semantic content (e.g. 'the dog had to be put down', 'she won the lottery jackpot') either matched that conveyed by prosody, as in the two congruent conditions, or conflicted with it, as in the two incongruent conditions. Sentences were approximately the same length, and of consistent style and format. A survey (n=20) determined perceived 'happiness' or 'sadness', and those sentences rated closest to the happy and sad ends of the rating scale were recorded by an experienced, male phonetician in happy and sad emotional intonation. Audiocassette recordings were digitised at 22 kHz/16 bits. Further background details on the generation of these stimuli can be found in Mitchell, Elliott, Barry, Cruttenden and Woodruff (2003). Stimuli were preferred in which relevant lexico-semantic information was presented across the whole sentence (as opposed to an isolated word in a 'carrier sentence'), to avoid interpretation difficulties caused by differences in stimulus onset asynchrony between lexico-semantic and prosodic elements (Wurm & Vakoch, 1996).

In each condition, twenty sentences were randomly selected from the pre-recorded suite, happy and sad scenario sentences being of equal number and randomly interspersed. The sentences within each condition were played one every 5 seconds, in a single block, via background noise attenuating headphones. In the congruent and incongruent attend semantic content conditions, participants' task was to concentrate on the emotion conveyed by semantic content and indicate whether they heard a sentence about a happy or sad scenario by pressing the H key (using their dominant hand) for happy sentences or S key for sad sentences. In the prosody conditions, participants' task was to concentrate on the emotion conveyed by tone of voice and indicate whether the speaker used a happy or sad tone of voice (same response keys as before). Participants were instructed not to wait until the end of a sentence to respond; they could indicate their response as soon as they had made their judgement. The order of the four blocks was fully counterbalanced across participants.

Comparable features of the colour and auditory emotional interference tasks included: (i) Both were block design paradigms, (ii) Mean phasic SCR responses to stimuli were collected in both, (iii) Both were presented via a standard PC, (iv) Both required a manual keyboard response via a single key, (v) Both tasks incorporate interference at the response selection/preparation stage (the irrelevant stimulus attribute clashed with the required response because the response preparation and interference processes share the verbal modality) (vi) The conditions within both tasks were counterbalanced (to avoid priming effects), (vii) The length of the conditions in each task was the same – 100s, and (viii) In both tasks, the task relevant and irrelevant dimensions are part of the same percept (e.g. unlike counting Stroop tasks; Bush, Whalen, Rosen, Junike, McInerney & Rauch, 1998).

Acquisition of Psychophysiological Data

Psychophysiological data was acquired in a noise and temperature controlled room (mean temperature 22.2°C ($\pm$1.78)). Prior to data collection participants' hands were washed with mild soap and water. Each participant's test session began with a 10-minute rest period during which participants' heart rate (HR) and SCR acclimatised to the environment. Further (3-minute) rest periods followed all conditions, to allow participant's HR and SCR to settle towards baseline. HR and SCR were recorded continuously using a BIOPAC MP-100 Remote Monitoring System (BIOPAC Systems Inc., California, U.S.A.). HR was obtained (in units of 'beats per minute' – BPM) by detecting the R wave from EKG electrodes placed on each wrist, with a reference electrode on the non-dominant forearm. SCR was measured (in μhos) using standard silver/silver chloride electrodes (8mm) filled with an isotonic conductive jelly (0.05 molar NaCl) and attached to the distal phalanx of the index and second fingers of the non-dominant hand. Importantly, doing so left the dominant hand free for behavioural responses. Participants were instructed to rest their hands immediately in front of the keyboard between stimuli, to minimise the movement and time required to make responses. A constant voltage of 0.5V was applied across the electrodes and the lower frequency response was set to DC, to acquire absolute skin conductance levels. Output from all electrodes was recorded on a standard PC using AcqKnowledge software (Biopac Systems – Goleta, California: http://biopac.com). During the acquisition of psychophysiological data, participants were asked to remain as still as possible, to avoid confounding these measurements.

Analyses

Behavioural Data:

For both tasks, the dependent behavioural measures were mean performance accuracy and mean reaction time accompanying correct responses. For the control task, data from each condition were entered into a paired samples t-test analysis to determine whether reaction times for the incongruent and congruent conditions differed significantly. For the auditory emotional interference task, data were entered into a 2 x 2 within-subjects ANOVA, comprising an 'emotion dimension' factor (lexico-semantic content vs. prosody) and a 'congruency' factor (congruent vs. incongruent). To compare the extent of congruency effects between tasks, mean reaction time from the congruent condition was subtracted from that in the incongruent, to create a single 'congruency effect size' measure for each task (collapsed across the lexico-semantic and prosodic conditions for the experimental task). Scores for the two tasks were then compared directly in paired t-tests. Parallel between-test comparisons of performance accuracy data were not possible due to ceiling effects on performance accuracy in the classic Stroop task.

Psychophysiological Data:

For both tasks, phasic skin conductance responses were isolated using AcqKnowledge to determine mean skin conductance levels during stimulus presentation, excluding that during inter stimulus intervals. For the analyses described, mean HR responses were similarly calculated over the period of stimulus presentation, excluding inter stimulus intervals. All SCR and HR data were corrected for baseline levels of responsivity; the mean measurement for the baseline immediately preceding each condition was subtracted from the mean measurement for each condition. If significant or repeated artefacts were observed in the traces during any condition, that data set was classified ‘missing data’. Minor or isolated artefacts were corrected by removing the affected part of the trace and connecting the disjointed end points.

Control Task0.0001 Tc -0.0001 Tsaseartefap (collapsed across the lexico-semantic and prosodic conditions for the experimental task). The scores for the two tasks were then compared directly in a paired t-test.

Results[1]

Behavioural Data (Table 2)

Classic Stroop Task: Participants took significantly longer to name the font colour in the incongruent condition than in the congruent condition, t(43) = 21.54, $p<0.001$, and participants were also less accurate in the incongruent condition than in the congruent, t(43) = 2.76, $p < 0.01$.

Auditory Emotional Interference Task: There was no effect of 'emotion dimension', F(1,43) = 0.30, p=0.59 on reaction time; reaction times were similar whether participants were responding to the emotion conveyed by lexico-semantic content or prosody. There was, however, a significant effect of congruency, F(1,43) = 5.19, $p<0.05$; participants took longer to respond to stimuli in which the emotions conveyed by lexico-semantic content and prosody conflicted. The interaction between emotion dimension and congruency was not significant, F(1,43) = 0.90, p=0.35; the effect of congruency on participants' reaction times was similar whether they were responding to lexico-semantic or prosodic emotion. The pattern of results was similar for performance accuracy: participants performed as accurately when attending to the emotion conveyed by lexico-semantic content as when attending to prosody (F(1,43) = 0.03, p=0.88), they performed less accurately in the incongruent condition than the congruent (F(1,43) = 8.15, $p<0.01$), and the interaction between emotion dimension and congruency was not significant (F(1,43) = 3.369, p=0.073).

Between-test comparisons of the effect of congruency on reaction time revealed that between-dimension incongruity in the classic Stroop task induced significantly

[1] Since they represented the majority gender in the participant sample, data from female participants was analysed on its own, to exclude possible partial confounds from male participant data. Analysis of female participants alone did not alter the significance patterns obtained at the behavioural, HR or SCR level. For clarity, corresponding descriptive and inferential statistics are therefore not presented here.

greater reaction time increases (from reaction times to congruent stimuli) than observed for the auditory emotional interference task ($t(43) = 7.64$, $p<0.01$).

Psychophysiological Data (Table 2)

Classic Stroop Task: Participants' SCR was significantly higher in the incongruent condition than in the congruent, $t(43) = 2.53$, $p<0.05$. In contrast, there was no difference between participants' HR in the congruent and incongruent conditions, $t(43) = 1.56$, $p=0.125$.

Auditory Emotional Interference Task: Preliminary comparisons of SCR and HR during the preceding baseline established that baseline SCR and HR measures did not differ significantly from one condition to another $F(1,43) = 2.10$, $p=0.12$ and $F(1,43) = 0.76$, $p=0.42$ respectively), thus avoiding the confounds suggested by Wilder's Law of Initial Values (1967). For the HR data analyses, the main effects of emotion dimension and congruency were not significant ($F(1,43) = 0.40$, $p=0.53$ and $F(1,43) = 0.45$, $p=0.51$ respectively). The interaction between emotion and congruency was not significant ($F(3,129) = 0.90$, $p=0.35$ either. For the SCR data, there was no main effect of emotion dimension ($F(1,43) = 1.94$, $p=0.17$). The main effect of congruency was significant, such that in the incongruent condition it was significantly higher than in the congruent condition ($F(1,43) = 4.11$, $p<0.05$). The interaction between emotion and congruency, was not significant ($F(3,129) = 2.94$, $p=0.09$).

Between-Task Differences: For the SCR data, 'congruency effect' scores were significantly higher for the classic Stroop task than for the auditory emotional interference task ($t(43) = 2.82$, $p<0.01$), but for the HR data, the difference between congruency effect scores for the tasks was not significant ($t(43) = 1.23$, $p=0.23$).

Table 2 about here

Discussion

Effect of Congruency:

As was the case with the classic Stroop task, inter-dimensional incongruency in the auditory emotional interference task caused a significant increase in SCR. Furthermore, this increase was similar whether participants were identifying the emotion conveyed by lexico-semantic content or that conveyed by prosody. The SCR pattern of response was paralleled by behavioural responses whereby participants were slower to respond in the incongruent condition than in the congruent. These results suggest that whether attending to lexico-semantic or prosodic emotion, the task-irrelevant dimension causes interference to the task-relevant dimension that is difficult to inhibit. Since SCR is believed to be related to arousal, the increased cognitive demand of the incongruent condition appears to have the psychophysiological consequence of increasing arousal, perhaps like the classic Stroop acting as a cognitive 'stressor', in a format potentially more useful to certain groups of researchers. According to Selye (1956), stress results when situational demands are perceived to exceed available coping resources, exemplified here by the presumable increased demand of the incongruent condition and concurrent detrimental effects on performance.

Whilst the interpretation of these results seems straightforward, the lack of previous consistency is somewhat puzzling. Some studies report a lack of congruency effects between lexico-semantic and prosodic emotion (Wurm & Vakoch, 1996; Wurm et al., 2001). However, these studies incorporated a task quite different to most other studies of this ilk, namely a lexical decision task about whether emotion 'words' at the end of carrier sentences, were real or non-words. In these stimuli, whilst prosodic information starts from the beginning of the sentence, emotional lexico-semantic information does not begin until near the end. Such a task also ignores any reverse influence of lexico-semantic emotion on prosodic emotion interpretation. The inter-dimensional interference in these tasks may not be strong enough to generate cognitive conflict. Bradley & Lang (2000) suggest that since new sensory information is added serially during presentation of auditory material spanning several seconds, physiological systems that respond to simple changes in the stimuli may be continuously active, making emotional signals difficult to detect. Nevertheless, like the shorter word-length stimuli presented by Grimshaw (1998) and Schirmer & Kotz (2003), it was still possible in the current study for participants to detect emotional signals and respond behaviourally and physiologically to the incongruity. As suggested by the authors themselves, using semantically and prosodically neutral stimuli too may be a further way to increase task demands such that congruency effects are more prevalent (Schirmer & Kotz, 2003), although doing so could mean that participants are no longer looking to identify the emotion conveyed by relevant dimension; by periodically comparing neutral to emotional information, they may in fact be thinking about the signal's intensity, a factor believed to relate to psychophysiological responses itself (Lang , Bradley & Cuthbert, 1998).

As to whether prosodic information interferes with the processing of lexico-semantic information more than vice versa, previous reports are perhaps, more contradictory. In the current study, there was a trend towards interaction between emotion and congruency, the effect of congruency on SCR appearing to be larger for prosody than for semantics, however, the difference did not reach statistical significance and neither did the corresponding behavioural data. As a 'low context' language, Kitayama and Ishii (2002) have proposed that English speakers rely less on contextual clues such as prosody, and more on lexico-semantic content. They found some support for the idea that the biggest interference experienced by English speakers was that of contrasting lexico-semantic evaluation on prosody evaluation. One could argue using a verbal response label, predisposes participants to lexico-semantic processing even when the task required them to focus on prosodic information. In contrast, Wurm and colleagues (2001) argue for the importance of contextual information such as emotional prosody, since it appears to constrain our interpretation of not just emotional words, but non-emotional too. Schirmer and Kotz (2003)' findings support Wurm et al's, but at the central nervous system level, their ERP data implied this pattern of response was limited to female participants. Whilst an imbalanced gender ratio could be considered a limitation of the current study, removal of male participants from the data set did not alter the significance patterns obtained at the behavioural or SCR level. In summary, currently available data provide evidence for the potential of significant cognitive conflict in both directions. Indeed, the current study provides behavioural indications of cognitive conflict whichever of the two dimensions is task-relevant. However, it extends previous literature by its original demonstrations of cognitive conflict during incongruent emotion via peripheral nervous system indicators. Despite our perpetual exposure to both sources of auditory emotion (and their episodic incongruence) and consequent potential equivalence in practice and automaticity, we should perhaps not discard the possibility that cognitive conflict resulting from incongruence of lexico-semantic and prosodic emotion may be flexible (Grimshaw, 1998).

Comparing the Tasks

Whilst the mean increase in SCR from baseline was 16.4% for the incongruent condition of the classic Stroop, the equivalent increase in SCR from baseline for the auditory emotional interference Stroop was only 7.5%. In addition, analysis of 'congruency effect sizes' revealed that the incongruent condition of the classic Stroop task caused a much larger increase in SCR from the congruent condition than did the experimental task. Similarly significant increases were observed in the accompanying behavioural (reaction time) data. The obvious explanation of this difference might simply be that the interference effect associated with the classic Stroop task is larger than that associated with the experimental task. One possible explanation for the apparent difference in effect size is that of different complexities in the possible number of conflict exemplars. Whilst the auditory emotional interference task incorporated two conflict exemplars (happy sentences spoken in a sad tone of voice, and sad sentences spoken in a happy tone of voice), the classic colour Stroop task incorporated several more (the word red in blue font, the word red in green font, the word blue in red font etc). However, the ratio of congruent:incongruent stimuli was the same in both tasks. One might also consider the time period of stimuli in each task a potential explanatory factor. Whereas in the classic colour Stroop task participants have to read a single word, in the auditory emotional interference paradigm, participants' behavioural and SCR responses reflect a longer period of stimulation. However, the purpose of the auditory emotional interference paradigm was that it represent the social situations in which we encounter conflicting emotion cues, as closely as possible (whilst retaining appropriate experimental testing conditions). In these situations we are more likely to encounter/have to process sentences than isolated words. Thus to assess cognitive conflict in these situations as naturalistically as possible, it was necessary to assess the responses to sentences rather than isolated words.

The smaller psychophysiological effect of the auditory emotional interference task might also relate to its 'everyday' nature. Whereas it is unlikely that most people have to distinguish between font colours and colour words in their daily lives, we are repeatedly faced with situations in which emotional communications are contradictory. In accordance with arguments raised in the introduction, it is possible that the practice we have of this skill means participants find the incongruent condition of the auditory emotional interference Stroop task less cognitively stressful than the colour task, hence the lower SCR. However, the statistical significance of the congruency factor in the auditory emotional interference Stroop task means that it is still an 'efficient' source of cognitive conflict. Whether its effects represent the same type of cognitive conflict as that induced by classic cognitive tasks like the Stroop test, remains somewhat unclear though. Further neuroimaging studies may prove particularly beneficial in this respect (see concluding thoughts). Similar neural mediators would suggest that these two sources of conflict represent a single process or mechanism.

HR Responses:

In contrast to the clear relationship between SCR and congruency, little effect of congruency was observed on HR. This result is perhaps not surprising considering the lack of predictable HR response to emotional stimuli of other modalities (Gendolla & Krusken, 2001; Gomez et al., 2004). Similarly, a reliable difference in HR between incongruent and comparison conditions is not always present in psychophysiological studies of the classic Stroop task (Silva & Leite, 2000). It seems that SCR changes are poorly correlated with HR changes (Gendolla & Krusken, 2001). This observation may well reflect the difference in innervation of these two psychophysiological indices, the eccrine sweat glands that contribute to SCR being innervated by the sympathetic division of the autonomic nervous system and the cardiovascular system being innervated by the sympathetic and parasympathetic divisions (in an antagonistic fashion). Indeed, emerging evidence suggests that different components of the autonomic nervous system may have different neural concomitants in the brain (Wittling, Block, Genzel & Schweiger, 1998a; 1998b).

Conclusions and Functional Neuroanatomic Implications

Whilst the effects may be subtler than those during standard cognitive paradigms, incongruence between dimensions of auditory emotion does cause notable cognitive conflict. Modulations of the effects of such emotional incongruity can be assessed from standard behavioural measures and certain psychophysiological measures. Early reports from neuroimaging studies of typical visual emotional Stroop tasks suggest some similarities between its mediation in the brain and that of classic colour Stroop tasks (Compton et al., 2003), particularly in the dorsolateral prefrontal cortex. However, the lack of inter-dimensional conflict in visual emotional Stroop tasks does not make them an ideal choice for studying the effects of emotional incongruence in the brain. In their fMRI investigation Schirmer, Zysett, Kotz and von Cramon (2003) conceptualised semantic-prosodic incongruence as a 'semantic anomaly'. Their discussion highlighted the association of inferior frontal gyrus activity with resolution of this incongruence, a region well known for its semantic functions (e.g. Matthews et al., 2003). However, the current study's multi-level demonstration of consequential cognitive conflict, raises the possibility that these early studies may be missing a potentially important focal point. Various cognitive paradigms have shown a strong link between mediation of cognitive conflict and recruitment of the ACC (Botvinick, Cohen & Carter, 2004). On the basis of evidence from the current study, future neuroimaging studies of emotional incongruence should perhaps consider this brain region in their a priori hypotheses. Bearing in mind that the effects of cognitive conflict can be expressed through alterations in autonomic nervous system activity, evidence of functional links between ACC activity and autonomic control (Critchley, 2002; Critchley et al., 2003; Matthews, Paulus, Simmons, Nelesen & Dimsdale, 2004) add weight to this suggestion. Whilst classic cognitive Stroop tasks tend to activate the posterior portion of the ACC, an emotion-based version such as this may be more likely to activate the anterior portion of the ACC (see Bush, Luu & Posner, 2000).

Acknowledgements

This study was performed whilst the author was a lecturer in the Department of Psychology at Keele University, Staffordshire, U.K.

References

Barrett, G. (2000). Clinical application of event-related potentials in dementing illness: issues and problems. International Journal of Psychophysiology, **37**, 49-53.

Borod, J. C. (1993). Cerebral mechanisms underlying facial, prosodic, and lexical emotional expression: A review of neuropsychological studies and methodological issues. Neuropsychology, **7**, 445-463.

Botvinick, M. M., Cohen, J. D., & Carter, C. S. (2004). Conflict monitoring and anterior cingulate cortex: An update. Trends in Cognitive Sciences, **8**, 539-546.

Bradley, M. M., & Lang, P. J., (2000). Affective reactions to acoustic stimuli. Psychophysiology, **37**, 204-215.

Brosgole, L., & Weisman, J. (1995). Mood recognition across the ages. International Journal of Neuroscience, **82**, 169-189.

Bush, G., Whalen, P. J., Rosen, B. R., Jenike, M. A., McInerney, S. C., & Rauch, S. L., (1998). The counting Stroop: an interference task specialized for functional neuroimaging--validation study with functional MRI. Human Brain Mapping, **6**, 270-282.

Bush, G., Luu, P., & Posner, M. I. (2000). Cognitive and emotional influences in anterior cingulate cortex. Trends in Cognitive Sciences, **4**, 215-222.

Campos, A., Marcos, J. L., & González, M. A. (1999). Relationship between properties of words and elicitation of skin conductance response. Psychological Reports, **85**, 1025-1030.

Carton, J. S., Kessler, E. A., & Pape, C. L. (1999). Nonverbal decoding skills and relationship well being in adults. Journal of Nonverbal Behavior, **23**, 91-100.

Compton, R. J., Banich, M. T., Mohanty, A., Milham, M. P., Herrington, J., Miller, G. A., et al. (2003). Paying attention to emotion: An fMRI investigation of cognitive and emotional Stroop tasks. Cognitive, Affective, & Behavioral Neuroscience, **3**, 81-96.

Critchley, H. D. (2002). Electrodermal responses: What happens in the brain? The Neuroscientist, **8**, 132-142.

Critchley, H. D., Mathias, C. J. Josephs, O., O'Doherty, J., Zanini, S., Dewar, B.-K., et al. (2003). Human cingulate cortex and autonomic control: Converging neuroimaging and clinical evidence. Brain, **126**, 2139-2152.

Dollins, A. B., Cestaro, V. L., & Pettit, D. J. (1998). Efficacy of repeated psychophysiological detection of deception testing. Journal of Forensic Science, **43**, 1016-1023.

Gaillard, A. W. (1988). Problems and paradigms in ERP research. Biological Psychology, **26**, 91-109.

Gendolla, G. H., & Krusken, J. (2001). The joint impact of mood state and task difficulty on cardiovascular and electrodermal reactivity in active coping. Psychophysiology, **38**, 548-556.

Gomez, P., Stahel, W. A., & Danuser, B. (2004). Respiratory responses during affective picture viewing. Biological Psychology, **67**, 359-373.

Grimshaw, G. M. (1998). Integration and interference in the cerebral hemispheres: Relations with hemispheric specialization. Brain and Cognition, **36**, 108-127.

Holland, P. C., & Gallagher, M. (2004). Amygdala-frontal interactions and reward expectancy. Current Opinions in Neurobiology, **14**, 148-155.

Hubert, W., & de Jong-Meyer, R. (1990). Psychophysiological response patterns to positive and negative film stimuli. Biological Psychology, **31**, 73-93.

Hugdahl, K. (1995). Psychophysiology: The Mind-Body Perspective. Massachusetts: Harvard University Press.

James, W. (1884). What is an emotion? Mind, **9**, 188-205.

Kitayama, S., & Ishii, K. (2002). Word and voice: Spontaneous attention to emotional utterances in two languages. Cognition and Emotion, **16**, 29-59.

Lang, P. J., Greenwald, M. K., Bradley, M. M., & Hamm, A. O. (1993). Looking at pictures: Affective, facial, visceral, and behavioural reactions. Psychophysiology, **30**, 261-273.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1998). Emotion, motivation, and anxiety: Brain mechanisms and psychophysiology. Biological Psychiatry, **44**, 1248-1263.

Macleod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. Psychological Bulletin, **109**, 163-204.

Macleod, C. M., & Dunbar, K. (1988). Training and Stroop-like interference: Evidence for a continuum of automaticity. Journal of Experimental Psychology: Learning, Memory and Cognition, **14**, 126-135.

Matthews, P.M., Adcock, J., Chen, Y., Fu, S., Devlin, J. T., Rushworth, M. F., et al. (2003). Towards understanding language organisation in the brain using fMRI. Human Brain Mapping, **18**, 239-247.

Matthews, S. C., Paulus, M. P., Simmons, A. N., Nelesen, R. A., & Dimsdale, J. E. (2004). Functional subdivisions within anterior cingulate cortex and their relationship to autonomic nervous system function. Neuroimage, **22**, 1151-1156.

McConnell, C. R. (2004). Interpersonal skills: What they are, how to improve them, and how to apply them. Health Care Management, **23**, 177-187.

Meadows M. -E., & Kaplan R. F. (1994). Dissociation of autonomic and subjective responses to emotional slides in right hemisphere damaged patients. Neuropsychologia, **32**, 847-856.

Melara, R. D., Rao, A., & Tong, Y. (2002). The duality of selection: Excitatory and inhibitory processes in auditory selective attention. Journal of Experimental Psychology: Human Perception and Performance, **28**, 279-306.

Mitchell, R., Elliott, R., Barry, M., Cruttenden, A., & Woodruff, P. W. R. (2003). The Neural Response to Emotional Prosody, as Revealed by Functional Magnetic Resonance Imaging. Neuropsychologia, **41**, 1410-1421.

Schirmer, A., & Kotz, S. A. (2003). ERP evidence for a sex-specific Stroop effect in emotional speech. Journal of Cognitive Neuroscience, **15**, 1135-1148.

Schirmer, A., Zysett, S., Kotz, S. A., & von Cramon, Y. (2004). Gender differences in the activation of inferior frontal cortex during emotional speech perception. Neuroimage, **21**, 1114-1123.

Selye, H. (1956). The stress of life. New York: McGraw-Hill.

Shanks, D. R., & Johnstone, T. (1999). Evaluating the relationship between explicit and implicit knowledge in a sequential reaction time task. Journal of Experimental Psychology: Learning, Memory and Cognition, **25**, 1435-1451.

Silva, F. T., & Leite, J. R. (2000). Physiological modifications and increase in state anxiety in volunteers submitted to the Stroop color-word interference test: A preliminary study. Physiology and Behavior, **70**, 113-118.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. Journal of Experimental Psychology, **28**, 643-662.

Trauner, D. A., Ballantyne, A., Chase, C., & Tallal, P. (1993). Comprehension and expression of affect in language-impaired children. Journal of Psycholinguistic Research, **22**, 445-452.

Tulen, J. H., Moleman, P., van Steenis, H. G., & Boosma, F. (1989). Characterization of stress reactions to the Stroop Color Word Test. Pharmacology, Biochemistry and Behavior, **32**, 9-15.

vanOyen Witvliet, C., & Vrana, S. R. (1995). Psychophysiological responses as indices of affective dimensions. Psychophysiology, **32**, 436-443.

Wilder, J. (1967). Stimulus and response: The law of initial value. Bristol, U.K.: Wright.

Wittling, W., Block, A., Genzel, S., & Schweiger, E. (1998a). Hemisphere asymmetry in parasympathetic control of the heart. Neuropsychologia, **36**, 461-468.

Wittling, W., Block, A., Schweiger, E., & Genzel, S. (1998b). Hemisphere asymmetry in sympathetic control of the human myocardium. Brain and Cognition, **38**, 17-35.

Wurm, L. H., & Vakoch, D. A. (1996). Dimensions of speech perception: Semantic associations in the affective lexicon. Cognition and Emotion, **10**, 409-423.

Wurm, L. H., Vakoch, D. A., Strasser, M. R., Calin-Jageman, R., & Ross, S. E. (2001). Speech perception and vocal expression of emotion. Cognition and Emotion, **15**, 831-852.

Table 1

Previous studies implementing paradigms that incorporated conflict between the emotions conveyed by lexico-semantic content and prosody.

Study	Purpose	Methods	Main Findings
Grimshaw (1998)	Examined interaction between lexico-semantic/prosodic information & hemispheric specialisation style (complementary specialization: process the 2 dimensions in opposite hemispheres vs. non-complementary: both dimensions processed in same hemisphere).	Dichotic listening task. Words ‘mad’, ‘sad’, ‘glad’, or ‘fad’ spoken in angry, sad, happy or neutral tone of voice. Participants attended to lexico-semantic content or tone of voice & indicated whether target information present in either ear or absent.	Lexico-semantic information interfered more with prosdic processing than vice versa, although prosodic information did produce significant interference. Failed to support hypothesis that corpus callosum shields each hemisphere from other to permit independent parallel processing – complementarity had no consequences for integration of lexico-semantic & prosodic information, or for their interference.
Kitayama & Ishii	Tested hypothesis that processing systems brought to bear on comprehension of	Positive & negative words spoken in ‘pleasant’ or ‘harsh’ tone of voice.	Interference effect by competing lexico-semantic evaluation in prosodic emotion

(2002)	emotional speech are attuned primarily to lexico-semantic evaluation in low-context culture & language (English), but are attuned primarily to prosodic emotion in high-context culture & language (Japanese).	Participants instructed to attend to 1 dimension whilst ignoring the other. Both dimensions tested in separate blocks. Self-paced task.	judgement stronger in English. Interference effect by competing prosodic emotion in lexico-semantic judgement stronger in Japanese.
Schirmer & Kotz (2003)	Compared interaction of emotional prosody & lexico-semantic valence during emotional comprehension in men & women.	ERP study. Positive, negative & neutral verbs spoken in happy, neutral & angry prosody. Participants instructed to attend to 1 dimension whilst ignoring the other. Both dimensions tested in separate blocks. ISI 3-3.2s.	Comparable behavioural effects between men & women: reactions to congruent stimuli faster & more accurate (more salient for lexico-semantic condition). ERP: smaller N400 amplitude for congruent stimuli, but only significant for lexico-semantic judgement & for female listeners.
Wurm & Vakoch (1996)	Affective lexicon explained in terms of 3 underlying dimensions: evaluation, activity & potency. Assessed importance of these	Pure emotion (disgusted, petrified & happy) & equivalent non-words inserted at end of single carrier phrase. Phrase read in happy,	Tone of voice did not influence decision times, nor did it interact with lexico-semantic category. Lexical decision times

	dimensions during speech perception.	petrified, disgusted, & neutral tones of voice. Participants made speeded lexical decision about target word.	significantly predicted by interaction between evaluation, activity & potency dimension weights.
Wurm et al. (2001)	To determine whether prosodic expression of emotion affected speed with which listeners could identify emotion words.	Same as for Wurm & Vakoch (1996).	Randomised presentation of tone of voice showed no effect of congruence. Blocked presentation of tone of voice did.

Table 2

Summary Descriptive Statistics

Task	Subtask	Condition	Accuracy (%)		Reaction Time (s)		Heart Rate* (beats per minute)		Skin Conductance Rate* (µmho)	
			Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation
Colour Stroop		Congruent	99.98	0.13	0.68	0.08	7.43	0.93	1.18	0.12
		Incongruent	98.36	3.90	1.28	0.17	9.30	1.12	1.86	0.26
Experimental Task	Attend Semantics	Congruent	91.48	8.11	2.38	0.26	0.52	4.01	0.75	0.72
		Incongruent	89.55	14.66	2.72	0.18	-0.65	3.00	0.79	.77
	Attend Prosody	Congruent	94.55	6.36	2.52	0.29	-0.65	3.69	0.55	0.57
		Incongruent	87.05	14.72	2.81	0.56	0.98	12.41	0.87	0.69

* Following correction of the data for baseline values